DOCUMENT RESUME

ED 116 060 CG 010 229

AUTHOR Block, Jack
TITLE Recognizing the Coherence of Personality.
PUB DATE Jul 75
NOTE 72p.

EDRS PRICE MF-$0.76 HC-$3.32 Plus Postage
DESCRIPTORS *Behavioral Science Research; Personality; *Personality Assessment; *Psychological Studies; *Social Psychology; State of the Art Reviews
IDENTIFIERS *Mischel (W)

ABSTRACT

In the last half dozen years, the field of personality psychology has been asked to reconsider certain premises that traditionally have guided conceptualization and research. The most influential questioner and critic has been Mischel. The essential criticism, advanced and buttressed in a number of ways, is an empirical one: namely, the research evidence accruing over the years supports only weakly or not at all the assumption by pshchologists that traits or dispositions importantly govern behavior. Mischel develops this conclusion in a widely-read review of the personality assessment literature. Mischel's 1968 conclusions have become widespread and blandly accepted. This paper offers an "assessment of Mischel's assessment of the state of personality assessment," and indicates some of the ways in which his negative evaluation can be countered. This essay reads the research evidence differently and, in addition, introduces some recent pertinent findings that permit a different structuring of the accomplishments and deficiencies characterizing personality research. (Author/HMV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED116060

Partial Draft
July 1975

RECOGNIZING THE COHERENCE OF PERSONALITY

Jack Block

University of California, Berkeley

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

I. Personality Assessment Under Siege

In the last half dozen years or so, the field of personality psychology has been asked to reconsider certain premises that traditionally have guided conceptualization and research. The most influential questioner and critic of personology has been Mischel (1968; 1969; 1972; 1973) but also expressing their concern have been Petersen (1965; 1968), Bem (1972), and Fiske (1973; 1974) among others.

The critique of personality psychology has been broadly put and has ranged widely. The esssential criticism, advanced and buttressed in a number of ways, is an empirical one: namely, the research evidence accruing over the years supports only weakly or not at all the assumption by personologists that traits or dispositions importantly govern behavior. Mischel develops this conclusion in a widely-read review of the personality assessment literature (1968). He then goes on to suggest that the disappointing accomplishments of personality psychology after so many years of extensive effort may well have a larger implication, that the "paradigm" (Kuhn, 1962) traditionally employed in personality

CE 010 229

psychology is fundamentally inadequate and should be replaced by newer conceptualizations rising above the limitations set by earlier, now demonstrably unproductive assumptions.

For Mischel, the immediate issue confronting personality psychology is not necessarily a conceptual one but derives instead from undeniable empirical insufficiencies. "The initial assumptions of trait-state theory were logical, inherently plausible, and also consistent with common sense and intuitive impressions of personality. Their real limitation turned out to be empirical--they simply have not been supported adequately." (Mischel, 1968, p. 147) In Mischel's view, this inadequate empirical support cannot be ascribed solely to faulty (but in principle, remediable) research methodology; rather, he is inclined to believe this "basic dilemma of evidence" calls into question the very paradigm personologists have been employing.

This sense of paradigmatic crisis has proven challenging; other psychologists now have responded to various issues underlying the several arguments brought together by Mischel. Thus, Bowers (1973) considers more closely some of the implicit assumptions underlying alternative approaches to psychology; Averill (1973) and Wiggins (1974) discuss the logical status of the concept of "disposition"; Alker (1972) and Wachtel (1973a; 1973b) elaborate various conceptual complications that might explain disappointing empirical relationships.

Much of what is said in these several articles I would myself endorse. However, these responses all seem to start

from (and thereby implicitly accept) Mischel's assessment of the state of personality assessment. Excepting only Craik's brief demurral (1969) regarding the evidential basis for Mischel's highly negative evaluation of assessment accomplishments, the discussions to date of Mischel's diagnosis have adopted, without question, his 1968 interpretation of the empirical symptoms. Further, although in later articles, Mischel has importantly qualified (and thereby softened) certain implications of his position, he has continued to reference his 1968 book to support his critical evaluation of the state of personality psychology. Of perhaps largest significance, Mischel's 1968 conclusions have become widespread and are cited frequently and with bland acceptance in texts and by journal editors as the received view of the field. It therefore becomes important and even required to offer an assessment of Mischel's assessment of the state of personality assessment, to indicate some of the ways in which his negative evaluation can be countered. The present essay reads the research evidence differently and, in addition, introduces some recent pertinent findings that permit a different structuring of the accomplishments and deficiencies characterizing personality research.

II. Mischel's Empirical Conclusions

The three main inter-related conclusions of Mischel assert:

1. There is little strong evidence for personality

consistency and much strong evidence for the specificity of behavior. Across situations and across time, the consistencies posited by trait- and psychodynamically-oriented psychologists only seldom have been found. See Mischel, 1968, Chapter 2.

2. Traits and states are largely constructs of the observer. The available evidence indicates that, more than a little, the personality traits and states observed by psychologists are the constructions of the observers rather than being strongly related to independent information about the subject's actual behavior. See Mischel, 1968, Chapter 3.

3. Inferences from traits and states are not predictively useful. Efforts to use trait and psychodynamic concepts in psychological prediction have not consistently demonstrated useful increments over easier, cheaper procedures for prediction. See Mischel, 1968, Chapter 5.

The first of Mischel's conclusions is the crucial one and accordingly is the primary concern of this essay. If this first summary evaluation by Mischel can be replaced by another perspective, then his second and third corollary conclusions also change in their significance because, as will be seen, the evidence to be brought forward for the existence of appreciable personality consistency also includes evidence that the traits and states viewed by observers derive in fundamental ways from qualities residing within the individuals observed and, in addition, are predictively useful.

III. A Worldly View of the State of Psychological Research

Before responding to Mischel's pessimistic conclusions with more optimistic perspectives--indeed, in order to respond--some prefatory remarks are necessary about such matters as the quality of contemporary psychological research and the inferences to be drawn from this research.

The quality of psychological research. In order to support a positive view of the possibilities residing within the trait and psychodynamic view, it is necessary to berate with little mercy much of the research in personality psychology. This is an ironic position, reluctantly but, I believe, realistically advanced. I hope I will not be viewed as putting myself forward as a lordly, non-pareil judge or critic. I simply (simply?) wish to assert that an intelligent, informed, vigilant evaluation of the recorded personality research will quickly reveal that perhaps 90% of the studies are methodologically inadequate, without conceptual implication, and even foolish. Without recognition of the many kinds of deficiencies characterizing personality (and other psychological) research, it is impossible to make sense of the "litter-ature." Paradoxically, critical evaluation will permit a meaningful view of the field; uncritical acceptance of current empiricism will lead to a nihilism. These strong words may seem unduly confessional for someone identified with the science of personality psychology, but such realizations, if indeed justified, are a necessary basis for the improvement and advancement of the field.

What is the basis for this indictment of so much personality research? Only some general remarks are feasible here, addressed to problems in operationalizing concepts, problems in formulating hypotheses, and problems of methodology and data analysis.

1. Many concepts have not been well-represented by their operational indicators. Psychologists have been extraordinarily casual and even irresponsible in developing measures to represent concepts. In coordinating a concept with a measure, it is incumbent upon the psychologist to justify, via construct validation procedures, the propriety of this assigned correspondence (Block, 1968a). Too often, psychologists will award an implicative, even flashy name to a particular measure, without supporting convergent and discriminant evidence for the label being employed. Whereupon, given the sociology of psychology, if the measure is a convenient one, other psychologists will employ it in a mountainous mass of studies of molehill significance. So, unless a trait or psychodynamic concept has been provided with a fair and supported operational translation, research involving these concepts is without implication.

As a case in point, consider the study by Twain (1957) who employed measures of height, weight, attitude toward the Chinese and attitude toward Germans, all as measures of impulsivity. He offered no supporting rationale or evidence for these posited operational definitions. Twain concluded there was little utility to the concept of impulsivity because

his various measures had little correlation with each other. Of what import is his finding?

2. Behavioral hypotheses derived from concepts should reflect the complexity and the implications of the concept being studied. Concepts nave to be thought about; they often have complicated or contingent or interactive implications which should be but are not respected in psychological research. Thus, psychological conceptualizations of the development of an ethical sense posit the gradual progression of the individual through a number of moral states (Loevinger, 1966; Kohlberg, 1964). Ethical consistency, according to these conceptions, cannot be manifested until certain ego or moral stages have been achieved. With this recognition in mind, consider the well-known Hartshorne and May (1928) studies of grade-school children, which frequently have been interpreted as evidence for the inconsistency of moral conduct. Clearly, these children were too young to have achieved the character stages required before consistency in moral behavior properly could be expected. Supporting this conceptual argument is the finding in the Hartshorne and May research that moral consistency increases with age. Moreover, Shuttleworth (in Hartshorne, May, & Shuttleworth, 1930) reports a finding that consistent subjects tend to be honest while inconsistent children tended to be dishonest. Given the theoretical formulations regarding moral behavior as further supported by close data analysis, it follows that the expectation by Hartshorne and May, which was accepted by Mischel, that behavior in various moral situations would inter-

correlate very highly was poorly based. The rejection of an unwarranted hypothesis obviously carries no implication for the usefulness of a concept or for the separate likelihood of supporting a warranted hypothesis.

3. The reliability of many of the measures employed in personality research and the power of the research designs employed is often poor, unnecessarily so. It makes no sense to use measures so unreliable that subsequent intercorrelations among measures are constrained to be close to zero. It also is unreasonable to interpret research, by oneself or by others, when the reliabilities of the measures used is unknown or not taken into account. One cannot know, in this latter circumstance, whether low intercorrelations are interpretable as such or are first ascribable to the poor quality of the measures involved. As I have previously noted (Block, 1963; 1964), although it is often impossible to have a precise reliability coefficient in hand for evaluating the attenuating effects of unreliability, a knowledgeable psychologist can, from prior or other knowledge, establish reasonable or constructive bounds for the reliabilities of the measures being evluated. These reliability estimates or guesses should be used. It is not enough to know in the abstract of the attenuating effects of reliability; rather, this recognition should be explicitly and specifically applied in the evaluation of each and every study. By taking attenuation effects into account, the research evaluator will develop a wiser perspective on the relationships involved.

Similarly, the embarrassingly frequent use of inefficient designs, as noted for example by Cohen (1962; 1969) also attenuates the possibility of discerning relationships. Psychologists will often employ research designs almost guaranteed to obscure relationships that might be residing in the data. Further, they are often misinformed regarding the way chance operates in statistical decision-making situations and biased to expect relationships where they should not exist (Tversky & Kahneman, 1971). When power-deficient research designs are further conjoined with unreliable measures and then evaluated by psychologists with rash expectations, it should not be surprising if strong and reproducible relationships seem to emerge only rarely.

Still a further example of the usage of analytical procedures that cannot, in principle, test the hypothesis supposedly under investigation is to be found in the number of studies that have relied upon the easy method of factor analysis to test for the existence of a general or consistency factor. Conventionally, a host of measures presumably equivalent or related as indicators or manifestations of a trait or psychodynamic concept are factored but almost without exception, the properties of the factor analytic method and of typical rotational procedures have gone unrecognized. As a consequence, results dictated necessarily by the method have been offered as substantive findings. For example, results are usually reported based upon the varimax method of rotation. But, in the varimax method, "a general factor is precluded" (italics

added) by the simplicity constraint on each factor (Harman, 1967, p. 304). Obviously, a method that prevents a general factor from emerging is the wrong tool with which to search for consistency.[1]

The Inferences to be Drawn from Contemporary Personality Research. If one accepts current research evidence in personality largely at face value, it is understandable why disappointment in empirical achievements to date has become widespread. Whereupon the search for alternative, potentially more successful approaches is launched. The positive suggestions by Mischel draw upon the empirical state of affairs but in a way that should be recognized as inductive rather than deductive, as tenable rather than required. The logical form of Mischel's reasoning seems to be as follows: behavior appears inconsistent and highly specific to the situation involved. If personality were indeed inconsistent, then behavior would appear inconsistent; therefore, personality is indeed inconsistent.

This reasoning is inductively supportable; it is not a deduction, however, since it reasons from a consequent to an antecedent. To put the point in another, related way, a failure to reject the null hypothesis of personality inconsistency does not prove the null hypothesis.

An equally sound induction from the empirical literature can reason that if personality research is poorly executed, then personality will appear inconsistent. If it is the case that personality research frequently is poorly operationalized and that appropriate hypotheses often are tested in insensitive ways,

what can one conclude regarding the usefulness of the trait and psychodynamic approaches from the variety of weak and certainly erratic relationships generated over the years? Rather little, I would suggest. Only after appropriate testing of the trait and psychodynamic approaches has failed would abandonment of this conceptual view seem due.

In the meanwhile, and as a supplementary consideration, it may well be that the current dismal assessment of the personality literature depends too heavily on the poor "batting average" our sloppy empiricism has attained. Home runs have been averaged with strike-outs, and clearly there have been many of the latter. But some people know how to play ball and others do not. What if the home runs are hit by competent, resourceful athletes while the strike outs come from the blind and the infirm? Surely, it is not elitist or beyond practical possibility to suggest that the "batting average" of personality psychology must be evaluated more closely, to see whether a pattern of quality or relevance can be said to characterize the order-implying findings reported, in contrast to a pattern of methodological insufficiency or conceptual irrelevance characterizing the results suggesting behavioral incoherence. I further note the important principle that the existence of only one good instance of research success is enough to prove the argument regarding possibility; a host of failures of bad research does not outweigh this possibility, once demonstrated.

There are many kinds of problems in applying critical stan-

dards to research. It is difficult and often impossible from published accounts for a reader to separate the empirical wheat from the empirical chaff. There can be selective, projective reading of the evidence, seeing virtue in the research reporting preferred relationships and seeing deficiencies in the research reporting relationships unassimilable to one's conceptual position. Great responsibility is required in the "empirical integrator" (Underwood, 1957); megalomania is to be avoided. But the job is worth attempting (indeed, is there an alternative?) and has the promise of discerning an order or structure that cannot otherwise be seen.

IV. Mischel's Appraisal of the Consistency and Specificity Literature

The first argument of Mischel is that the postulated and widely accepted notions using traits or psychodynamic "genotypes" to account for the consistencies of behavior beg the question--in his opinion, there is little empirical evidence of these consistencies for the trait or psychodynamic approaches to explain.

In my own view, there is decent and sometimes even impressive empirical support for the trait and psychodynamic approaches. Moreover, I believe there is a structure to the evidence. The pattern of where the findings are strongly positive and where they are discouraging has strong and even ironic implications for understanding the problems besetting personality psychology. Later sections will bring forward some of this evidence and its implications; here, some remarks are offered

on why Mischel's negative evaluation of the literature on behavioral consistency need not be discouraging to personologists.

A detailed, point by point countering evaluation will not be delivered because it is not required. What is required is an indication of how the thrust of his assessment can be parried by certain recognitions or alternative perspectives.

To begin, note that Mischel's summary regarding the consistency and specificity of behavior relating to personality variables is quite brief, 5507 words and less than 16 pages (Mischel, 1968, pp. 20-36). Within these few pages, he touches on attitudes toward authority and peers, moral behavior, sexual identification, dependency and aggression, rigidity and tolerance for ambiguity, cognitive avoidance, conditionability, moderator variables, and the temporal instability of personality (612 words per topic). Mischel did not intend these pages to be viewed as a sufficient or close evaluation of the available literature in the tradition of the lengthy reviews that appear in the Psychological Bulletin. Rather, these pages served Mischel as a vehicle to illustrate his perspectives and conclusions on the issues involved albeit in a highly distilled form. Obviously, Mischel's conclusion, whatever its degree of correctness, cannot be truly supported by so brief, selective, and undetailed a literature presentation. Instead, as Bowers (1973) has noted, we must view Mischel's conclusions as deriving from certain larger premises and orientations he holds regarding psychology.

Consider now the very first example Mischel offers of a

representative personality disposition for which the evidence of cross-situational generality is disappointing, the issue of attitudes toward authority and peers (Mischel, 1968, pp. 21-23). In Mischel's view, the trait and psychodynamic approaches have assumed the existence of highly generalized reactions to the various authority figures encountered in life. I myself believe, along with Wachtel (1973), that at least the psychodynamic viewpoint regarding authority relationships and transference is far more complicated in its position and should not be operationally reduced to the empirical hypothesis that all measures of attitudes toward authority should interrelate. Surely, this hypothesis is wrong; there are patterns of identification. One may like one's father but not want to be like him; one may hate one's father but want very much to be the powerful person he appears to be. And so on. Of more immediate import, however, is the cogency of the Burwen and Campbell (1957) research, the only reference cited by Mischel to evaluate this certainly disputable hypothesis. The onerous detail that follows in describing this study is a necessary burden if we are to evaluate this research rather than simply accept its summary conclusions.

The Burwen and Campbell study employed 73 officers and 82 enlisted men from 17 bomber crews. Burwen and Campbell report "perfunctory compliance and occasional humorous sabotage of the test purpose" (p. 24), with the result that 10 to 15% of the answer sheets were discarded. Compulsory interviews were held with 57 of the subjects. "Particularly for enlisted personnel, this setting created a guarded, deferential attitude

that was difficult to overcome during the one hour period involved" (Burwen and Campbell, 1957, p. 25). Clearly, questions can be raised about the meaning or credibility of data collected in such testing circumstances.

All the measures were constructed for the purposes of the Burwen and Campbell study and so there was no prior evidence for their construct validity. Further, there has been no subsequent research to demonstrate the validity of the measures employed. The measures included the following:

1. A single interviewer rated the covert attitudes of each interviewed subject as favorable/unfavorable, on a 5-point scale, toward his father, toward his siblings, toward his superior officers, and toward his peers. No reliabilities are presented for these four scores but the authors suggest the squared correlation coefficient of these measures with other test variables can provide a minimum estimate of the reliabilities. These figures were .16, .26, .31, and .08, respectively.

2. Stories written by the subjects in response to 8 TAT-like pictures were objectively evaluated to provide favorable/unfavorable scores, on a 5-point scale, for each subject with respect to his attitudes toward Symbolic Authority and toward Symbolic Peers. The usage of older persons and persons of higher military rank in the TAT stories was assumed to be symbolic of authority figures; the usage of peer-age persons as persons of lower rank was interpreted as symbolic of peers or subordinates. The reliabilities of these two scores were .55

and .53, respectively. Global clinical ratings of these two dimensions were also used; their lower bound reliabilities were .04 and .13, respectively.

3. The subjects employed a 30-item adjective check list to describe first their father, then an immediate superior (boss) and finally, a fellow worker. Scoring was in terms of the number of favorable terms employed in each description. Internal consistency reliabilities for these measures were .24, .34, and .55, respectively.

4. The subjects wrote two or three sentences to characterize 25 photos projected before them. Twenty of the photos were of individuals over 45 years of age; five photos were of high school age persons. The characterizations written by each subject for each photo were rated as favorable, intermediate, or unfavorable. The characterizations of individuals over the age of 45 were interpreted as an index of attitude towards Symbolic Authority; the characterizations of high school students were interpreted as a measure of attitude toward Symbolic Peers. Reliabilities of these two scores were .46 and .29, respectively.

5. Subjects completed an autobiographical inventory which contained free response items and checklist items. The inventory was scored using subjective ratings of the free response items and an a priori keying of the checklist. Scores were developed to reflect the favorableness/unfavorableness of the subject toward his father, toward a boss, and toward his peers. The reliabilities of these three scores were .56, .56, and .55, respectively.

6. An attitude survey was administered to the subjects to determine their attitudes, as measured on a 5-point scale, toward the Air Force generally, toward superior officers, and toward members of the subject's bomber crew. Attitudes toward the Air Force generally and toward superior officers were combined to form a score reflecting attitude toward Symbolic Authority. The subject's attitudes toward members of his own bomber crew was taken as a measure of his attitude toward Peers. The reliabilities of these two scores were .59 and .73, respectively.

7. A sociometric questionnaire was administered which defined five hypothetical situations for which the subject was asked to pick a companion from among the members of his bomber crew. An enlisted man who picked his aircraft commander relatively often was considered to be favorable toward authority; an enlisted man who tended to nominate other enlisted persons was considered favorable towards his peers. Officers could not develop scores for their attitude toward authority but those officers who tended to nominate other officers were scored as favorable toward peers. The reliabilities for the enlisted men were .75 and .90, respectively; for the officers, .84.

The measures employed in the Burwen and Campbell study have been described in detail to provide context for the results issuing from this test of the assumption of a generalized attitude toward authority. Clearly, the testing situation was an unfortunate one. Clearly, reliabilities were often very low and precluded the finding of appreciable cross-measure relationships.

Clearly, too, the operational translations of the authority concept were not validated and are highly contestable.

What were the results? There is a significant instrument component (i.e., method variance) underlying the measures. But, of 169 correlations not involving the same instrument, 143 (or 85%!) are positive, a highly significant finding suggesting some degree of commonality rising above method variance in the host of measures employed. Although the Burwen and Campbell averaged correlations reported by Mischel in his Table 2 are of modest size, there is very great variability about these averages and some correlations are quite high. Moreover, attenuation effects were not taken into account in evaluating these correlations.

Overall, it is possible to argue that this study, given its unfortunate testing circumstances and the poor reliabilities associated with its ad hoc controversial measures, has issued results that are astonishingly supportive of the notion of trait consistency. My own preference, however, is to set this study aside as simply irrelevant to the issues supposedly being studied. In my view, the "totally negative" conclusion advanced by Burwen and Campbell (1957, p. 31) and reiterated by Mischel (1968, p. 23) is premature because of the many methodological and operationalizing problems afflicting the study.

It is readily possible to challenge or to counter other aspects of Mischel's appraisal of the literature on personality consistency and specificity. Thus, Burton summarizes his well-known reanalysis of the Hartshorne and May studies as follows: "The conclusion to draw from these analyses is not greatly different from

that made by Hartshorne and May, but the emphasis on lack of relation between tests is removed." (Burton, 1963, p. 492). Mischel prefers to emphasize the first portion of Burton's conclusion; for many other personologists, it is the second portion of Burton's conclusion that is noteworthy because it represents a fundamental change, based upon a better analysis, in the previously received interpretation of the Hartshorne and May investigations.

As another instance of specificity, Mischel points to an absence of correlation among children's behaviors conceptualized as different facets of dependency. However, these various facets include behaviors (e.g., positive attention-seeking and negative attention-seeking) that perhaps for psychodynamic reasons, but also and certainly for definitional, zero-sum reasons can be expected to be related negatively in at least some of the children studied. Thus, when the base rate or time spent in both positive attention-seeking and negative attention-seeking is <u>low</u>, it is possible for the behaviors to covary positively. However, if the base rate of one of these behaviors is high in a time-sampled context, then the frequency of the other behavior <u>must</u> be low, a logically entailed negative correlation. Evaluation of a sample of children with different patterns of dependency behavior will issue a mish-mash correlational result.

Bandura's (1960) report that delinquent boys are prudently nonaggressive when with their harshly punishing parents, but are aggressive with school peers is cited by Mischel as still another bit of evidence for behavioral specificity. But surely and easily, such behavior can be viewed by the trait or psychodynamic per-

sonologist as consistent with the character structure variously labeled as sado-masochistic, dominant-submissive, or opportunistic (Loevinger & Wessler, 1970) wherein the individual has learned to place himself within a pecking order, deferring to those above him and behaving aggressively with those of equal or inferior rank.

The three variables, resistance to temptation, guilt feelings, and knowledge of moral standards do not correlate especially with each other, a finding Mischel cites as evidence for specificity. But where is the theory that says these variables should relate positively? The individual difference variable, resistance to temptation, by its very usage acknowledges that information as to what is moral or what is forbidden is insufficient to guarantee resistance; doing the forbidden causes remorse in some individuals, but not in others.

Clearly, Mischel has some unacknowledged presumptions about trait or psychodynamic conceptualizations that cause him to anticipate certain behaviors should be related. This anticipation, when disappointed, permits him to conclude trait or psychodynamic variables have failed. But this gloominess need not influence the mood of other personologists with different anticipations.

These last ripostes at Mischel's evaluation have been quickly made; their collective aim is to indicate that conceptual issues and understandings become complex very quickly. An evaluation of empirical insufficiencies must be attentive to attendant conceptual complexities, if it is to be relevant and therefore compelling. But perhaps the best response to the

view that empirical findings do not square with conceptual expectations is to attempt to show where and when they do and where and why they do not.

To this effort, we now turn.

V. Consistency Within and Between Personality Data Domains

As Cattell noted years ago (Cattell, 1957; 1973), it is useful to distinguish among three kinds of personality data--O-data, S-data, and T-data, through which traits or dispositions can be operationalized.

1. O-data are data derived from observers' evaluations of individuals leading more or less natural lives. Generally, these data take the form of personality ratings.[2]

2. S-data are data derived from the self-observations of individuals regarding their behavior, feelings, and characteristics. Self-ratings and responses to personality inventories or questionnaires exemplify these kinds of data.[3]

3. T-data are data derived from standardized, objective, more of less artificial test or laboratory situations wherein selected, specific, readily identified or enumerated behaviors are focused upon, unbeknownst to the participating subject, as indicators of particular personality variables. For example, the time a child takes to make a decision in circumstances involving response uncertainty has been employed as an indicator of the child's reflectivity-impulsivity (Kagan, Rosman, Day, Albert, & Phillips, 1964).

I wish to call attention to the nature and extent of the

disposition or trait consistency existing within and between these three data domains. It will be my contention, to be supported by some recent or unrecognized research results, that:

1. Well-done O-data studies demonstrate undeniable and impressive personality consistency and continuity reside within the individuals being studied.

2. S-data studies using carefully constructed personality inventories also show indisputable and appreciable personality coherence and stability within the individuals studied.

3. There are strong relationships between the dispositional qualities of individuals as studied via O-data and as evaluated using S-data.

4. The evidence for personality consistency as derived from studies using T-data is extremely erratic, sometimes positive, but often not.

5. As a corollary of the inconsistency manifested by T-data, it follows that the relationship between T-data on the one hand, and either O-data or S-data on the other hand must also be uneven.

The above five recognitions can place a different perspective on the state of personality assessment. In particular, the deficiencies or irregularities existing within the T-data domain carry implications and directives of great consequence which I will only be able to begin to discuss. But first to bring together some of the evidence bearing on these summary assertions.

<u>There is appreciable personality consistency and continuity as studied within the O-data domain</u>. Ratings of personality

have, over the years, earned an unfortunate reputation. They are costly, require the rater to live with uncertainty, and have been generated often in confounded, biased, subjective, unreliable ways. But costs must be judged relative to goals, observer-judges can accept the necessity of decision, and there is no reason why prior practice must control rather than inform subsequent efforts to improve the quality of O-data. Complicated and burdensome though they may be, observer-evaluations of personality can be employed in fully rigorous ways meeting the usual scientific criteria of data reproducibility within any data set and independence among the data sets subsequently related (Block, 1961, Chapter 3). What are some such studies and what results issue from them?[4]

1. The well-known longitudinal studies at Berkeley over the years collected an enormous amount of naturalistic information on a large number of boys and girls, men and women. My book, Lives Through Time (Block, 1971), integrates this material in an account, to date, of the personality characteristics and personality development of the individuals under study. The research design imposed upon the archival material relied heavily on O-data carefully developed.

For each subject, the naturalistic information available for the junior high school years was collected as one data set. Separately, naturalistic information for the senior high school years was assembled to construct a second data set. And finally, information available from an extensive interview of the subject during his or her fourth decade became a third data set.

These three data sets were strictly independent, no data in one set being carried over into another.

The material for a particular subject at a particular age was evaluated by (usually) three clinical psychologists, each functioning independently. No psychologist evaluated a subject at more than one age and, moreover the combinations of psychologists judging each particular age were permuted extensively, using the large pool of psychologists-judges available, to prevent the possible introduction of systematic judge effects. Psychologists expressed their descriptions of formulations of each subject using the California Q-Set procedure (Block, 1961). The California Q-Set provides a basic and reasonably well-established set of variables for the psychodynamic descriptions of personality, conjoined with an improved rating methodology that prevents the intrusion of extraneous and obfuscating differences between judges in their categorizing tendencies. Interjudge agreement in their CQ-formulations was generally acceptable and consequently, for each subject at each time period, the several CQ-formulations were arithmetically averaged. Thus, a consensually-based and reproducible composite CQ-description was available for each subject, at each of three time periods--during junior high school, during senior high school, and as an adult. Further extensive information regarding the research design, the procedures employed, and the quality controls applied as the data were developed is available in Lives Through Time (Block, 1971). What is important to note here is the independence among the judges in their personality formulations;

the large number of judges employed in permuted combinations; the use of a judge only once for a particular subject; the use of the Q-sort procedure so that personality formulations issuing from a variety of psychologist-judges readily could be made comparable and usable; and the use of three partitioned, strictly independent time-separated sets of largely naturalistic data, different in important respects for each subject.

Given the care and logic underlying these O-data, it would appear difficult to explain away substantial and conceptually expectable relationships empirically found to exist between time periods. Such relationships cannot be attributed to the effects of common data, common subjects, or the subtle influences of stereotypes. Rather, such relationships if obtained can be most readily (and perhaps only) understood in terms of enduring qualities within the subjects studied, qualities that were manifest in diverse ways, but were recognizable in their implications by experienced clinical psychologists.

And just what, in fact, are the findings? Over the 3 year period from junior high school to senior high school, 59% (67/114) of the personality variables (CQ-items) characterizing the male sample display consistency significant at the .001 level or better (correlations of at least .35, uncorrected for attenuation). For the female sample, the corresponding figure is 57% (65/114). Over the period from senior high school to the mid-thirties, an interval averaging close to 20 years, 28% (25/90) of the CQ-items show consistency significant at the .001 level or better (correlations of at least .35, uncorrected

for attenuation). Within the female sample, the corresponding figure is 30% (27/90). The correlations, uncorrected, range as high as .70 and .61 for the junior high school-senior high school and senior high school-adult intervals, respectively. If the unreliabilities of the individual CQ-items are allowed for, many of the correlations exceed .6 or .7.[5]

To sketchily sample the psychological nature of the temporal consistencies observed, the CQ-item, "Is a genuinely dependable and responsible person" correlates .58 in the male sample from JHS to SHS and .53 from SHS to adulthood; the CQ-item, "Tends toward undercontrol of needs and impulses, unable to delay gratification" correlates .57 from JHS to SHS and .59 from SHS to adulthood. The CQ-item, "Is self-defeating" correlates .50 from JHS to SHS and .42 from SHS to adulthood. The CQ-item, "Enjoys aesthetic impressions, is aesthetically reactive" correlates .35 from JHS to SHS and .58 from SHS to adulthood.

Within the female sample, the CQ-item, "Basically submissive" correlates .50 from JHS to SHS and .46 from SHS to adulthood. The CQ-item, "Emphasizes being with others, gregarious" correlates .39 from JHS to SHS and .43 from SHS to adulthood. The CQ-item, "Tends to be rebellious and non-conforming" correlates .48 from JHS to SHS and .49 from SHS to adulthood. The CQ-item, "Is concerned with philosophical problems, e.g. religion, values, the meaning of life, etc." correlates .45 from JHS to SHS and .42 from SHS to adulthood. All of these correlations are uncorrected for attenuation; many more could be cited. The

consistencies relate to personality qualities, not simply or or primarily to intellectual or cognitive characteristics where Mischel does acknowledge consistency exists.

When it is further recognized that over these extended periods of time, appreciable character change and transformation must have been involved in many of the individuals studied and that an overall correlation coefficient is a most inefficient and easily misled index of relationship, it seems to me that these O-data results provide altogether impressive evidence of personality consistency. The findings of identifiable personality transformation become even more striking, in my view, if the analytical approach becomes more differentiated. Using these O-data, I derived via inverse factor analysis a number of homogeneous types of personality development. Certain types of adult character structure can be identified with astonishing fidelity in early adolescence. For details as to these findings, many concomitant relationships, and the reationale for this approach, the reader will have to consult Lives Through Time (Block, 1971, Chapters 7, 8, 9, and 10).

2. In an ongoing longitudinal study of ego and cognitive development being conducted at the University of California, Berkeley by my wife, Jeanne H. Block, and myself, we have been studying something more than 100 children during their fourth, fifth, sixth, and eighth years of life. Various experimental procedures designed to reflect particular personality dimensions have been administered to these children during each testing year. My focus here, however, is on the O-data we have collected on these children.

The California Child Q-Set, an age-appropriate modification of the California Q-Set (Block, 1961; 1971), was used to develop personality characterizations of each child. The California Child Q-Set consists of 100 widely ranging, personality-relevant items that are ordered, using a forced-choice method, by a trained judge to express the judge's characterization of the personality of the child. The judges employed to characterize each child were his or her nursery school teachers, three teachers for more than half of the children and two different teachers for the remaining children. In judging a child, each teacher worked completely independently of the other teachers and based her personality formulations on 5 to 9 months of observation of the child's behavior in the nursery school setting for 3 hours each day. Thus, each child was well known by each judge, and the salient, consistent qualities of each child's personality had an opportunity to become manifest. All five nursery school teachers received training and calibration in using the Q-set before contributing their evaluations of these children. With the completion of the many Q-sorts, for each child the two or three Q-descriptions independently formulated by his or her teachers were averaged, resulting in one composite personality characterization for that particular year. This procedure was followed during the child's fourth year and also during his fifth year. The five nursery school teachers contributing their personality formulations during the fourth year were an entirely different set from the five nursery school teachers contributing characteri-

zations during the fifth year.

These data are most simply compared normatively rather than ipsatively by evaluating for each Q-item the orderings of the children developed independently and a year apart. Within the convergent-discriminant framework (Campbell and Fiske, 1959), these across-time correlations can be viewed as evidence of convergent validity of the trait ratings since the usage of different and independent sets of judges employing different information at two different times in effect results in different "methods" of measurement of the personality variables being studied. The discriminant validity of each variable can be evaluated by noting whether it correlates higher at another time with differently-named variables than it does with its correspondingly-named variable.

With these conceptions of convergent and discriminant validity in mind, consider our results. For the 100 Q-items, the average across-time correlation, calculated via the z-transformation, is .48, uncorrected for attenuation. The three lowest across-time item correlations are -.08, .14, and .18. The three highest item correlations are .70, .70, and .70. This average level of correspondence, by contemporary standards of psychological research, is rather high. With respect to discriminant validation, for 45% of the Q-items the correlation between the Q-item as rated during the fourth year and that same Q-item as independently rated during the fifth year was higher than any of the correlations of that fourth year rating with the 99 other Q-ratings for the fifth year. Thus,

not only appreciable convergence, but also appreciable discrimination characterizes these O-data. The absolute independence between these personality characterizations developed a year apart means that the relationships observed derive from qualities and consistencies within the children being studied and cannot be attributed to the personal constructs or attribution tendencies of the judges offering their ratings.

Some additional analyses of these data seeking to improve their convergent and discriminant validity are instructive, pointing up a moral and perhaps a solution for a problem that goes generally unrecognized. I hope you will tolerate what at first appears a digression.

In inspecting failures of discriminant validation, we observed that many of these "misses" were "good misses," misses that made a psychological sense. As one example, the Q-item, "Has rapid shifts in mood; emotionally labile" correlates .51 from the fourth to the fifth year, but the correlation of this item as measured during the fourth year with another Q-item as measured during the fifth year is an even higher .58, a nominal failure of discriminant validity. However, when it is noted that the second Q-item is "Overreacts to minor frustrations; easily irritated," the higher correlation between these two different items can be seen to derive from their conceptual connectedness or equivalence. Such "good misses" are not recognized properly within the usual convergent and discriminant framework--a miss seems to be as good as a mile. But a close miss is better than a far miss and indeed, it can be instructive

to study close misses to see why they were not "hits."

The frequency and nature of these "good misses" suggested to us that we had been unfair to teacher-raters contributing their Q-formulations. Apparently, a number of the Q-items were, when applied to young children, redundant or beyond the ability of the judges to discriminate. We were exceeding the psychological resolution capacity of the judges. So, it seemed sensible to bring together these correlated, unreliably different Q-items in order to develop broader and better variables. Accordingly, we factored and varimaxed the fifth year Q-items and decided that 12 factors could be said to encompass the data. Beyond these 12 factors, we had only a few doublets and residual items. For each of the 12 factors, factor scores were derived by standard scoring the several factor-loading Q-items and then averaging the standard scores, for both the fifth year Q-data on which the factor analyses were based and for the entirely independent fourth year Q-data. Thus, the pattern of the factor results from the fifth year data determined how scores were derived from the fourth year data, an arrangement that from one perspective (e.g. canomical correlation) is less than optimal when maximal correspondence between the two sets of data is sought. The advantage of applying the factor scoring or weighting arrangement based upon the fifth year data to the fourth year data as well is that there is absolutely no capitalization on chance to bolster unfairly the relationships between age levels and that fully equivalent sets of variables exist at both ages.

We are now to the point of this statistical excursion. It will be recalled that the set of 100 Q-items displayed 45% discriminant validity. What is the discriminant and convergent validity of the set of 12 factors, intermediate level variables presumably better meshed to the discrimination possibilities available from our personality characterizations?

For the twelve factors, identically measured in the fourth and fifth years, the discrimination validity is 100%! Each factor from the fourth year displays its highest correlation with its corresponding factor as measured in the fifth year. The mean across-time correlation or convergent validity for these 12 factors is .56, a figure most would agree is impressive.

Again for illustrative purposes and to breathe life into these summary figures, the factor, "Compliance," as rated during the fourth year correlates ,72, uncorrected for attenuation, with "Compliance" as independently rated during the fifth year. The factors, "Undercontrol," "Resilience," and "Empathic Relatedness" as rated during the fourth year correlated, respectively, .71, .46, and .64 with their corresponding factors during the fifth year. These figures are also uncorrected for attenuation. Although these data could readily be improved upon, if simply by using more judges, these convergent validities are already and impressively high. Since they are based upon fully independent sets of data, they are difficult to ascribe to artifact or to the workings of constructs solipsistically held by the observers involved.

Rather, the results indicate that even within young children, recognizable and perduring qualities of personality have been formed.

The larger implication of these analyses is that the usual convergent validity-discriminant validity approach does not distinguish between failures of discriminant validation that are conceptually readily assimilable. For practical and conceptual reasons, the sets of variables we employ may be inappropriate and hence unable to manifest convergent and discriminant validity. Active conceptual and analytic effort, perhaps judiciously using procedures like factor analysis, can help move us toward a set of personality variables that is discriminating, reliable, and interesting.

Summarizing now our presentation regarding O-data, it has been shown that good quality and independently established O-data displays appreciable and encouraging convergent and discriminant validity. Other illustrations beyond those cited here can be culled from the literature (e.g., Gormly & Edelberg, 1974). My strong impression is that unpublished data in the archives of the Institute of Personality Assessment Research will also support the general findings advanced here regarding O-data. It should also be recognized that the data reported, although of decent quality, could well be improved upon with the consequence that the convergent and discriminatory relationships reported can be expected to become better.

Why is it that O-data functions in so orderly a way? The reasons are several and with large implication. We hold off

this discussion until after presenting our perspective and information on the convergent and discriminant validity of S-data and T-data.

There is appreciable personality consistency and continuity, as studied within the domain of S-data. Self-report questionnaires and personality inventories have a long history in psychology (Goldberg, 1971) and widespread usage. The dominant inventory still is the Minnesota Multiphasic Personality Inventory (MMPI) but the California Psychological Inventory (CPI), which derives substantially from the MMPI has had extensive usage as well. Rising in popularity in recent years are Cattell's 16 Personality Factor Questionnaire (the 16 PF) and the Personality Research Form of Jackson. A variety of studies have shown that the dimensions measured by one inventory or questionnaire usually can be measured impressively well by alternative inventories. For a summary of these studies together with a demonstration of the interchangeability of the CPI and the 16 PF, see the report by Campbell and Chun (1975). These findings that alternative and independent inventory-based measurements of personality dimensions are highly related are a first and extensive indication that individual differences in personality, as quantified by S-data, are consistent. Let me add some additional evidence, of a different kind, to this essential conclusion.

1. The CPI has had a long and productive history since its introduction by Gough (1957; 1964). As typically employed, 18 scales are scored although the 480 item pool can be employed

to score a number of other personality dimensions as well. The 18 scales generally scored are labelled as follows: Dominance, Capacity for Status, Sociability, Social Participativeness, Sense of Well Being, Responsibility, Socialization, Self Control, Tolerance, Good Impression, Communality, Academic Achievement via Conformance, Academic Achievement via Independence, Intellectual Efficiency, Psychological-Mindedness, Flexibility, and Femininity. The extended meaning of these scales, the basis of their derivation, their reliabilities, validities and associated relationships may be found in other sources (Gough, 1964; Megargee, 1972).

The CPI was administered to adult subjects in the Berkeley longitudinal studies on two separate occasions, 10 years apart. It would be difficult to argue that the subjects remembered ten years later their specific responses to the 480 items. Separating the two sexes within each of the two longitudinally-studied samples, four independent samples can be identified: men (N = 39) administered the CPI at ages 38 and 48; men (N = 59) administered the CPI at ages 31 and 41; women (N = 43) administered the CPI at ages 38 and 48; and women (N = 78) administered the CPI at ages 31 and 41. What is the convergent and discriminant validity of the 18 CPI scales over the ten-year period involved?[6]

For the four separate samples, the discriminant validity (meaning an inventory scale correlates higher with itself ten years later than it does with any other scale ten years later) are 89% (16 of 18 scales), 100% (18 of 18 scales), 89% (16 of 18 scales) and 100% (18 out of 18 scales)! The

very few failures of discriminant validity are by small amounts and make obvious psychological sense. The mean convergent validities are .68, .70, .72, and .73! Recognizing the effects of unreliability and of genuine personality change over the ten years involved, it would appear that these figures could hardly be higher. I also suggest that these findings are probably of general applicability in the S-domain; they are not unique to the CPI. The evidence on inventory interchangeability indicates that other substantial inventories would have done as well, had they been employed.

2. Many of the subjects in the longitudinally studied sample had been administered the WILTD Questionnaire (WILTD) during their junior high school and senior high school years in the 1930's. The WILTD questionnaire consisted of 50 questions regarding the subject's preferences and tendences in a number of life situations. No special rationale underlies the questions employed and the wording of the questions as formulated in these early days leaves much to be desired. For the purposes of Lives Through Time and as reported therein, I factor analyzed the WILTD questionnaires, separately and combined for the sexes and the time periods involved, with the result that two primary and overriding factors seem to be present. The one factor was labelled, somewhat vaguely, as "bland socialization." It was measured by only ten items and need not concern us here. The second factor, however, impressed me as a clear expression of "overcontrol." I developed scores for each subject by simply summing across the 17 items loading on this

factor for each subject. To convey quickly a sense of the items involved in this scale, here are two examples: "Can you stick to disagreeable work for a long time though no one makes you do it?" (scored for a True response) and "Do you get angry easily?" (scored for a False response).

For the purposes of the present paper, I thought to correlate the WILTD Overcontrol scores derived during the high school years with the CPI Ego Control and Self-Control scores derived from an administration of the CPI about 25 years later. The CPI Ego Control scale was constructed on the basis of criterion groups; the CPI Self-Control scale is a rational scale subsequently improved by internal consistency analyses. In any event, the measures being related were independently constructed and based upon data widely separated in time.

In the male sample, from junior high school to senior high school, the WILTD Overcontrol scales correlated .48. From senior high school to adulthood, the WILTD Overcontrol score correlated .52 with the CPI Ego Control scale and .50 with the CPI Self-Control scale. In the female sample, from junior high school to senior high school, the WILTD Overcontrol scales correlated .66. From senior high school to adulthood the WILTD Overcontrol score correlated .53 with the ego control scale and .43 with the self-control scale. None of these figures allow for attenuation due to unreliability. Considering the nonoptimal nature of the measures involved and the time span of a quarter century, the findings of these appreciable and conceptually required correlations is further firm evidence of

personality continuity in the S-data domain.

3. There are strong relationships between the qualities of individuals as studied via O-data and as evaluated using S-data. It has already been shown that O-data developed to describe subjects longitudinally studied display good personality consistency over time and that S-data collected longitudinally also reveal impressive personality continuity. Now, it remains to see whether these two data domains are strongly or at least sufficiently related, as they must be.

The group of individuals studied continues to be the sample from Lives Through Time, for whom Q-composites and CPI protocols exist. It is not entirely clear just what the best way of connecting the personality Q-ratings to the CPI protocols may be. An orthodox multivariate statistician might suggest canonical correlation or multiple regression techniques but these methods capitalize on chance, require larger sample sizes than psychologists usually have available, and provide results in a form usually not psychologically conveyable. I prefer, at least for the present purpose, a simple, ostensive and therefore readily understandable method. Specifically, what are the particular personality ratings significantly associated with the various CPI scales?[7] If these ratings are numerously and appropriately correlated with the CPI scales, it will be clear enough that the two data domains are related.

When the personality ratings characterizing the subjects during adulthood are related to the CPI scales administered at about the same time, the significant correlations observed across

the O- and S-data domains are both plentiful and psychologically relevant. These results are too voluminous to report here but an indication of the strength and conceptual validity of these across-domain associations earlier was presented in Chapters 8 and 9 of Lives Through Time where the many CPI scales significantly associated with a variety of rating-defined personality types are listed.

In the present paper, for dramaturgical reasons, I elect to report the connections between the CPIs administered when the subjects were in their mid-thirties and the personality ratings formulated to characterize the personalities of the subjects during adolescence, some 20 or 25 years earlier. I also report the connection between a questionnaire or S-data measure developed during adolescence with personality ratings formulated a generation later, when the subjects were in adulthood. Because of the absolute independence of the data domains and the great time spans involved, because of the many attenuating factors that operated, and because of the characterological changes that must have been present, correlations having statistical size and making psychological sense should be especially persuasive evidence for an essential coherence of personality.

In relating the CPI taken during adulthood to earlier personality ratings formulated in adolescence, again a profusion of statistically significant findings was observed. For economy of presentation, only the results surrounding two CPI scales will be reported. The two scales, the Dominance Scale and the

Socialization Scale, are central scales of the CPI; they have been carefully developed; and they were designed to measure quite different psychological dimensions. For our male sample, the correlation between these two scales was -.01; for the female sample, the correlation was -.09. Tables 1 through 4 present the Q-item rating correlates from both the junior high school and senior high school periods with the Dominance and Socialization Scales of the CPI taken 20 to 25 years later, for both the male and female samples.

Insert Tables 1 through 4 about here

I suggest that the reader who peruses the numerous correlates in these tables will recognize and will not dispute the existence of constellations of personal qualities that accord well with the generally-held meanings of dominance and socialization. Many of the correlations, although low, even so serve to augment the interpretation and implications of these concepts. Certainly, there are some differences between the sexes in the personality precursors of these dimensions as later measured. Overall, however, considering the many obstacles to discernment of relationship affecting these analyses, it seems fair to conclude that rich and required connections exist between these personality ratings and the CPI scales studied.

In relating questionnaire or S-data from adolescence to

personality ratings or C-data developed during adulthood, the only available questionnaire scale score was the WILTD Overcontrol scale, earlier described. Tables 5 and 6 report the Q-items from the personality formulations of the subjects as adults that correlate significantly with the WILTD questionnaire measure of Overcontrol, administered when the subjects were in senior high school.

Insert Table 5 and 6 about here

Again, I suggest the tabled relationships demonstrate a strong correspondence between Overcontrol as measured via questionnaire in adolescence and rating-based personality characteristics of subjects, evaluated in their mid-thirties. As has been noted earlier, the manifestations of ego-control are different in males than in females because of differences in the prescriptive and proscriptive properties of sex roles (J. H. Block, 1973). Because of this recognition, better questionnaire measurement of ego-control is available when sex-specific inventory scales can be employed (Block, 1965). But even though this desirable approach was not feasible within the present data constraints, the findings testify to an enduring congruence between questionnaire and rating evaluations of overcontrol.

Recognizing the less-than-optimal nature of the data and measures being employed but recognizing too the strict separation

of the data between domains and across time, I believe there are grounds for encouragement and even a sense of security about certain principles and practices of personality psychology and personality assessment. It should also be noted that the findings just reported are by no means unique. With respect to the first two dimensions of the Minnesota Multiphasic Personality Inventory (MMPI), dimensions that are different from the CPI Dominance and Socialization dimensions here evaluated, O-data or personality rating correlates have been identified in five different samples (Block, 1965, Tables 19 through 28). Again, the connections between the O- and S-domains were plentiful and concordant. And finally, O-data or personality rating correlates of broad arrays of MMPI, CPI, and Strong Vocational Interest Inventory scales can be found in now old reports (Block & Bailey, 1955; Block & Petersen, 1955; Block & Gough, 1955). It would appear fair to conclude that the O- and S-domains have been linked in ways that, although improvable, are already quite substantial. No prestidigitation is required to achieve the results reported, simply the straight-forward but careful application of procedures so well-known as to be prosaic. Reasoning from past accomplishments, there is little reason to doubt that well-based and well-quantified O-data will continue to be strongly related to well-developed S-data scales.

4. The evidence for personality consistency as derived from studies using T-data is extremely erratic, sometimes positive but often not. This assertion should require little documentation since in large measure, the currently-held despondent views of personality consistency derive from such evidence. Repeatedly, investigators have observed that putatively equivalent or related measures in the T-data domain do not manifest their conceptually-required correspondence. Thus, Cole (1974) concludes there is little empirical support for the characteristic of "curiosity" as a behavioral disposition operating in different, supposedly curiosity-evoking situations. Chown (1959) in her evaluation of the concept of "rigidity" found little evidence of coherence among a variety of purported rigidity measures. Measures of "reflectivity-impulsivity" (Kagan, Rosman, Day, Albert, & Phillips, 1964), "motor-inhibition (Maccoby, Dowley, Hagen, & Degerman, 1965), and "delay of gratification" (Mischel, 1961) should, for conceptual reasons, be linked together but they are not (Shipman, 1971). Many of the studies cited by Mischel (1968) further exemplify the frequent failure of T-domain measures to interrelate as, conceptually, is to be expected. At will, one can wander through the pages of personality journals and find further instance after further instance of the absence of expected correlations among T-measures. It is because this point has been and can be documented so extensively that I elect not to make the case, in any detail, again here.

5. It follows as a corollary, therefore, of the erratic

relationships among T-measures, that the relationship between T-data on the one hand and either O-data or S-data on the other will also be uneven. Again, it is necessary only to exemplify rather than to document this conclusion. For convenience, I illustrate the problem by some data from the ongoing study by my wife and myself previously mentioned wherein we were interested in measuring "delay of gratification." One of the measures employed was a modification of the delay of gratification procedure earlier developed by Block & Martin (1955). In this experiment, the subject child worked for M & M candies and was permitted to accumulate as many M & M's as desired before stopping to eat and enjoy any. However, once having stopped to partake of the pleasures of sweetness, the child could not resume work to acquire more candy. Thus, a child presumably able to delay gratification could acquire many candies before stopping; a child presumably unable to delay gratification would acquire only a few candies before stopping. A second experimental procedure designed to tap the child's ability to delay gratification involved the child's reaction to a gaily-wrapped package identified as a present for him. The present, contents unknown, was shown to him and then ostentatiously set to the side by the experimenter who directed the child's attention to the completion of a jigsaw puzzle task. After four minutes, during which time the experimenter as required assisted the child to complete the puzzle, the child waited a further 90 seconds while the experimenter busied herself. During all this time, the package identified to the child as a present was in the child's

sight. At the end of the 90 seconds, the child was told he could have the present, if he had not already by then taken it. The child's delay time before taking the present during the 90 second interval constituted the score of interest, to represent delay of gratification. Both the candy acquisition experiment and the gift delay experiment were administered to the children at ages 3.5. and 4.5.

From the one age to the other, the correlation between candy acquisition scores was .24 for the boy sample and .30 for the girl sample. For the gift delay procedure, the across-time correlations were .23 for the boys and .03 for the girls. The correlations between the two procedures at 3.5 were .01 and -.29, for the boys and girls, respectively; and at 4.5, the correlations were -.07 for the boys and -.08 for the girls.

Of greater interest for the present purposes, however, are the correlates between these T-domain measures of delay of gratification and the O-domain personality ratings previously described. For the candy acquisition scores, at both ages 3.5 and 4.5, there were fewer significant personality correlates for either boys or girls than would have been expected on the basis of chance. The specific CCQ-item, "Is unable to delay gratification," correlated .09 and .11 (non-significant and in the wrong direction) with candy acquisition scores for the boys and girls respectively at age 3.5; at age 4.5, the correlations were, respectively, .20 and .11 for boys and for girls.

For the gift delay time scores, however, at both ages 3.5 and 4.5, there were many and conceptually congruent correlates

with the CCQ personality ratings. The specific CCQ-item, "Is unable to delay gratification," correlated -.43 in the sample of boys and -.50 in the sample of girls at age 3.5; at age 4.5, the corresponding correlations were -.30 and -.35 for the boys and girls, respectively. For illustrative purposes, Table 7 presents all the CCQ-item correlates of gift delay time for the boys and girls at age 3.5. The pattern of correlates at age 4.5, when the procedure was repeated, is not quite so strong although it is by no means weak.

Insert Table 7 About Here

Although there are some interesting and perhaps suggestive discrepancies between boys and girls in their respective patterns of CCQ-correlates with gift delay time, overall there appears to be good correspondence. A richly elaborated picture of the boy and of the girl who is unable to delay gratification is to be found in Table 7, a constellation of findings that suggests the gift delay time procedure is indeed "getting at" the concept intended. But why did not the candy acquisition procedure, also carefully designed and previously used, fail to generate the correlates needed to support its aspired-to validity? We do not really know, although we have some conjectures on the matter. For the moment, however, the only point requiring recognition is that this kind of anomaly, of erratic relationships between O-data and T-data, arises often and, since

we have shown O-data can function impressively well, the fault must lie with the insufficiencies of T-data.

Having shown that within the domains of O- and S-personality data, given good methodology, indisputably strong relationships exist and that within the domain of T-personality data, the evidence for lawfulness and coherence is far more difficult to attain, it is now incumbent upon us to consider why this pattern of law and disorder exists and what strategies are likely to extend the realm of coherence so as to include as well the domain of T-data.

Reference Notes

1. Bandura, A. Relationship of family patterns to child behavior disorders. Progress Report, U.S.P.H.S. Research Grant M-1734, Stanford University, 1960.
2. Campbell, J., & Chun, K-T. Comparison of the California Psychological Inventory and the 16 PF. Unpublished manuscript. Institute of Social Research, University of Michigan, 1975.
3. Wiggins, J. S. In defense of traits. Vancouver, B. C.: University of British Columbia, Unpublished Report, 1974.

References

Alker, H. A. Is personality situationally specific or intrapsychically consistent? Journal of Personality, 1972, 40, 1-16.

Averill, J. R. The dis-position of psychological dispositions. Journal of Experimental Research in Personality, 1973, 6, 275-282.

Bem, D. J. Constructing cross-situational consistencies in behavior: Some thoughts on Alker's critique of Mischel. Journal of Personality, 1972, 40, 17-26.

Block, J. The Q-sort method in personality assessment and and psychiatric research. Springfield, Illinois: Charles C. Thomas, 1961.

Block, J. The equivalence of measures and the correction for attenuation. Psychological Bulletin, 1963, 60, 152-156.

Block, J. Recognizing attenuation effects in the strategy of research. Psychological Bulletin, 1964, 62, 214-216.

Block, J. The challenge of response sets. New York: Appleton-Century-Crofts, 1965.

Block, J. Personality measurement. International Encyclopedia of the Social Sciences. The Macmillan Company, 1968, Volume 12, 30-37.

Block, J. Lives through time. Berkeley, California: Bancroft Books, 1971.

References (Continued)

Block, J., & Bailey, D. Q-sort item analyses of a number of MMPI scales. Technical Memorandum OERL-TM-55-7. Officer Education Research Laboratory, Maxwell Air Force Base, Montgomery, Alabama, May 1955.

Block, J., & Gough, H. G. An application of the Q-sort technique to the California Psychological Inventory. Technical Memorandum OERL-TM-55-8. Officer Education Research Laboratory, Maxwell Air Force Base, Montgomery, Alabama, May 1955.

Block, J., & Petersen, P. Q-sort item analyses of a number of Strong Vocational Interest Inventory Scales. Technical Memorandum OERL-TM-55-9. Officer Education Research Laboratory, Maxwell Air Force Base, Montgomery, Alabama, May 1955.

Block, J. H. Conceptions of sex roles. American Psychologist, 1973, 28, 512-526.

Block, J. H., & Martin, B. Predicting the behavior of children under frustration. Journal of Abnormal and Social Psychology, 1955, 51, 281-285.

Bowers, K. B. Situationism in psychology: An analysis and critique. Psychological Review, 1973, 80, 307-336.

Burton, R. V. Generality of honesty reconsidered. Psychological Review, 1963, 70, 481-499.

Burwen, L. S., & Campbell, D. T. The generality of attitudes toward authority and nonauthority figures. Journal of Abnormal and Social Psychology, 1957, 54, 24-31.

References (Continued)

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Cattell, R. B. Personality and motivation structure and measurement. New York: World Book Company, 1957.

Cattell, R. B. Personality and mood by questionnaire. San Francisco: Jossey-Bass, 1973.

Chown, S. Rigidity: A flexible concept. Psychological Bulletin, 1959, 56, 195-223.

Cohen, J. The statistical power of abnormal-social psychological research: A review. Journal of Abnormal and Social Psychology, 1962, 65, 145-153.

Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1969.

Coie, J. An evaluation of the cross-situational stability of children's curiosity. Journal of Personality, 1974, 42, 93-116.

Craik, K. C. Personality unvanquished. Contemporary Psychology, 1969, 14, 147-148.

Fiske, D. W. Can a personality construct be validated empirically? Psychological Bulletin, 1973, 80, 89-92.

Fiske, D. W. The limits for the conventional science of personality. Journal of Personality, 1974, 42, 1-11.

Goldberg, L. R. The proliferation of personality scales and inventories: An historical analysis. In P. McReynolds (Ed.) Advances in psychological assessment, Volume 2. Palo Alto, Calif.: Science and Behavior Books, 1971.

References (Continued)

Gough, H. G. Manual for the California psychological inventory. Palo Alto, Calif.: Consulting Psychologists' Press, 1957. Revised Edition, 1964.

Harman, H. H. Modern factor analysis. Chicago: University of Chicago Press, 1967.

Hartshorne, H., & May, M. A. Studies in the nature of character. (Volume 1) Studies in deceit. New York: Macmillan, 1928.

Hartshorne, H., May, M. A., & Shuttleworth, F. K. Studies in the nature of character. (Vol. 3) Studies in the organization of character. New York: Macmillan, 1930.

Kagan, J., Rossner, B. L., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. Psychological Monographs, 1964, 78 (1, Whole No. 578).

Kohlberg, L. Development of moral character and moral ideology. In M. Hoffman and L. W. Hoffman (Eds.) Review of child development research. Vol. 1. New York: Russell Sage Foundation, 1964. Pp. 383-431.

Kuhn, T. S. The structure of scientific revolutions. Chicago, Ill.: University of Chicago Press, 1962.

Loevinger, J. The meaning and measurement of ego development. American Psychologist, 1966, 21, 195-206.

Loevinger, J., & Wessler, R. Measuring ego development. Vol. 1. San Francisco: Jossey-Bass, 1970.

References (Continued)

Maccoby, E. E., Dowley, E. M., Hagen, J. W., & Degerman, R. Activity level and intellectual functioning in normal preschool children. Child Development, 1965, 36, 761-770.

Megargee, E. I. The California Psychological Inventory Handbook. San Francisco: Jossey-Bass, 1972.

Mischel, W. Preference for delayed reinforcement and social responsibility. Journal of Abnormal and Social Psychology, 1961, 63, 116-124.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Mischel, W. Continuity and change in personality. American Psychologist, 1969, 24, 1012-1018.

Mischel, W. Direct versus indirect personality assessment: Evidence and implications. Journal of Consulting and Clinical Psychology, 1972, 38, 319-324.

Mischel, W. Toward a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283.

Peterson, D. R. Scope and generality of verbally defined personality factors. Psychological Review, 1965, 72, 48-59.

Peterson, D. R. The clinical study of social behavior. New York: Appleton-Century-Crofts, 1968.

Shipman, V. C. Disadvantaged children and their first school experiences: Structure and development of cognitive competencies and styles prior to school entry. Princeton, N.J.: Educational Testing Service, 1971.

References (Continued)

Stephenson, W. The study of behavior. Chicago: University of Chicago Press, 1953.

Tversky, A., & Kahneman, D. Belief in the law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

Twain, D. C. Factor analysis of a particular aspect of behavioral control: Impulsivity. Journal of Clinical Psychology, 1957, 13, 133-136.

Underwood, B. J. Psychological research. New York: Appleton-Century-Crofts, 1957.

Wachtel, P. L. Psychodynamics, behavior therapy, and the implacable experimenter: An inquiry into the consistency of personality. Journal of Abnormal Psychology, 1973, 82, 324-334.

Waeftel, P. L. On fact, hunch, and stereotype. Journal of Abnormal Psychology, 1973, 82, 537-540.

Table 1

Personality Rating Correlates in Adolescence of CPI Dominance Scores Gathered When the Male Subjects Were in Their Mid-Thirties

JHS r	SHS r	Q-Item Content
.31b	.29b	Has a wide range of interests.
.22a	.23c	Is a talkative individual.
.30b	---	Appears to have a high degree of intellectual capacity.
-.24b	-.26b	Is uncomfortable with uncertainty and complexities.
-.28b	---	Basically submissive.
-.21c	---	Feels a lack of personal meaning in life.
-.21c	---	Tends toward overcontrol of needs and impulses.
.25b	---	Shows condescending behavior in relations with others.
.21c	.22c	Is turned to for advice and reassurance.
-.26b	-.25c	Gives up and withdraws in face of frustration/adversity.
-.26b	-.28b	Vulnerable to real or fancied threats.
-.38a	---	Reluctant to take definite action.
---	27b	Is facially and/or gesturally expressive.
-.24c	-.23c	Is basically distrustful of people/question motivations.
.27b	---	Genuinely values intellectual and cognitive matters.
.33a	.25c	Behaves in assertive fashion in interpersonal situations.
.25b	.27b	Is an interesting, arresting person.
-.21c	---	Concerned with body and adequacy of physiological function.
.29b	---	Has high aspiration level for self.
---	-.23c	Has clear-cut, internally consistent personality.
---	-.24c	Appears straightforward, forthright.

Block

Table 1 (Continued)

JHS r	SHS r	Q-Item Content
.21c	---	Is cheerful.
---	-.30b	Handles anxiety/conflicts by repression or disassociation.
---	.39a	Tends to proffer advice.
-.38a	-.38a	Is emotionally bland/has flattened effect.
.31b	---	Is verbally fluent/can express ideas well.
.21c	.25c	Is self-dramatizing/histrionic.
-.31b	-.25b	Does not vary role/relates to everyone in same way.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 2

Personality Rating Correlates in Adolescence of CPI Dominance Scores Gathered When the Female Subjects Were in Their Mid-Thirties

JHS r	SHS r	Q-Item Content
.29b	.39a	Is a talkative individual.
.24c	---	Appears to have a high degree of intellectual capacity.
---	-.33a	Is uncomfortable with uncertainty and complexities.
-.35a	---	Anxiety and tension find outlet in bodily symptoms.
-.32a	-.35a	Basically submissive.
.33a	.39a	Has rapid personal tempo.
---	-.31a	Arouses nurturant felling in others of both sexes.
---	-.35a	Feels a lack of personal meaning in life.
---	-.35a	Tends toward overcontrol of needs and impulses.
---	.24b	Is turned to for advice and reassurance.
-.25b	-.39a	Gives up and withdraws in face of frustration/adversity.
---	-.26b	Is calm, relaxed in manner.
-.34a	-.32a	Vulnerable to real or fancied threats.
---	.23c	Is moralistic.
-.25b	-.40a	Reluctant to take definite action.
.22c	.31b	Is facially and/or gesturally expressive.
-.28b	-.39a	Has brittle ego-defense system/maladaptive under stress.
-.21c	-.41a	Tends to feel guilty.
---	-.25b	Aloof/avoids close interpersonal relationships.
---	-.25b	Is basically distrustful of people/questions/motivations.
.22c	.41a	Behaves in assertive fashion in interpersonal situations.

Table 2 (Continued)

JHS r	SHS r	Q-Item Content
---	.28b	Emphasizes being with others/gregarious.
-.24b	---	Is self-defeating.
---	.24c	Responds to humor.
.35a	---	Is an interesting, arresting person.
-.29b	-.21c	Concerned with body and adequacy of physiological function.
---	.25b	Socially perceptive of wide range of interpersonal cues.
---	.22c	Pushes/stretches limits/sees what he can get away with.
---	.24b	Has high aspiration level for self.
.21c	---	Consciously unaware of self-concern/consistent personality.
-.28b	---	Projects own feelings and motivations onto others.
---	-.32a	Feels cheated and victimized by life.
---	-.22c	Ruminates and has persistent, pre-occupying thoughts.
---	-.35a	Handles anxiety/conflicts by repression or disassociation.
---	.21c	Is power oriented/values power in self and others.
---	.43a	Has social poise and presence.
---	.21c	Expresses hostile feelings directly.
---	.34a	Tends to proffer advice.
.22c	---	Values own independence and autonomy.
-.34a	-.31b	Is emotionally bland/has flattened effect.
---	.30b	Is verbally fluent/can express ideas well.
---	.21c	Is self-dramatizing/histrionic.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 3

Personality Rating Correlates in Adolescence of CPI Socialization Scores Gathered When the Male Subjects Were in Their Mid-Thirties

JHS r	SHS r	Q-Item Content
.53a	.48a	Is a genuinely dependable and responsible person.
-.22c	---	Is a talkative individual.
.41a	.29b	Behaves in giving way toward others.
.30b	.33a	Is fastidious.
.50a	.24c	Is protective of those close to him.
.43a	.47a	Behaves in a sympathetic or considerate manner.
.34a	---	Arouses nurturant feeling in others of both sexes.
-.21c	-.34a	Feels a lack of personal meaning in life.
-.23c	-.23c	Extrapunitive/tends to transfer or project blame.
.22c	.23c	Prides self on being objective, rational.
---	.35a	Tends toward overcontrol of needs and impulses.
.44a	.52a	Is productive/gets things done.
.38a	.33a	Tends to arouse liking and acceptance in people.
.23c	.30b	Is turned to for advice and reassurance.
.27b	---	Is satisfied with personal appearance.
.29b	.25c	Seems to be aware of the impression he makes on others.
.25b	.37a	Is calm, relaxed in manner.
-.23c	-.47a	Over-reactive to minor frustration/irritable.
.33a	---	Has warmth/is compassionate.
-.34a	-.29b	Is negativistic/tends to undermine/obstruct/sabotage.
-.49a	-.25c	Is guileful and deceitful, manipulative, opportunistic.
-.36a	---	Has hostility toward others.

Table 3 (Continued)

JHS r	SHS r	Q-Item Content
-.37a	-.24c	Has brittle ego-defense system/maladaptive under stress.
-.27b	---	Is basically distrustful of people/questions/motivations.
-.37a	-.48a	Is unpredictable and changeable in behavior and attitudes.
-.25b	-.44a	Undercontrol of needs, impulses.
-.40a	-.33b	Is self-defeating.
.29b	---	Has insight into own motives and behavior.
-.35a	-.38a	Tends to be rebellious and non-conforming.
-.22c	---	Judges self and others in conventional terms.
-.35a	-.25b	Pushes/stretches limits/sees what he can get away with.
-.25b	-.24c	Is self-indulgent.
-.24b	-.27b	Bothered by anything that can be construed as a demand.
---	.27b	Has high aspiration level for self.
.30b	.29b	Consciously unaware of self-concern/consistent personality.
.35a	---	Has clear-cut, internally consistent personality.
-.29b	---	Projects own feelings and motivations onto others.
.37a	---	Appears straightforward, forthright.
-.34a	-.30b	Feels cheated and victimized by life.
-.24c	---	Ruminates and has persistent, pre-occupying thoughts.
-.22c	---	Interested in members of opposite sex.
---	.23c	Is physically attractive/good-looking.
-.26b	-.44a	Has fluctuating moods.
---	.31b	Is cheerful.
-.30b	---	Interprets simple/clear-cut situations in complicated ways.
-.25b	---	Compares self to others.

Table 3 (Continued)

JHS r	SHS r	Q-Item Content
---	.28b	Has social poise and presence.
.23c	---	Behaves in a masculine or feminine style or manner.
---	-.39a	Expresses hostile feelings directly.
-.29b	-.23c	Is self-dramatizing/histrionic.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 4

Personality Rating Correlates in Adolescence of CPI Socialization Scores Gathered When the Female Subjects Were in Their Mid-Thirties

JHS r	SHS r	Q-Item Content
---	-.29b	Is critical, skeptical, not easily impressed.
.42a	.39a	Is a genuinely dependable and responsible person.
-.25b	---	Is a talkative individual.
.32a	---	Behaves in giving way toward others.
.30b	.39a	Is fastidious.
---	.27b	Is uncomfortable with uncertainty and complexities.
.25b	---	Is protective of those close to him.
.33a	.34a	Basically submissive.
---	-.24b	Is introspective.
.33a	.28b	Behaves in a sympathetic or considerate manner.
.37a	---	Arouses nurturant feeling in others of both sexes.
---	-.24c	Extrapunitive/tends to transfer or project blame.
.34a	.37a	Tends toward overcontrol of needs and impulses.
.37a	.41a	Is productive/gets things done.
.24c	.26b	Tends to arouse liking and acceptance in people.
---	.24c	Is satisfied with personal appearance.
---	.34a	Is calm, relaxed in manner.
-.21c	-.46a	Over-reactive to minor frustration/irritable.
.27b	---	Has warmth/is compassionate.
-.33a	-.40a	Is negativistic/tends to undermine/obstruct/sabotage.
-.30b	-.26b	Is guileful and deceitful, manipulative, opportunistic.
-.29b	-.31b	Has hostility toward others.

Table 4 (Continued)

JHS r	SHS r	Q-Item Content
---	-.25b	Thinks and associates ideas unusually.
---	-.31b	Is basically distrustful of people/questions/motivations.
-.40a	-.51a	Is unpredictable and changeable in behavior and attitudes.
-.39a	-.39a	Undercontrol of needs, impulses.
-.31a	-.36a	Is self-defeating.
---	-.21c	Is an interesting, arresting person.
-.27b	---	Enjoys sensuous experiences.
-.50a	-.53a	Tends to be rebellious and non-conforming.
.21c	---	Socially perceptive of wide range of interpersonal cues.
-.56a	-.34a	Pushes/stretches limits/sees what he can get away with.
-.35a	-.23c	Is self-indulgent.
-.21c	-.21c	Bothered by anything that can be construed as a demand.
---	-.23c	Perceives different contexts in sexual terms.
---	.38a	Consciously unaware of self-concern/consistent personality.
.24c	.37a	Has clear-cut, internally consistent personality.
---	-.21c	Projects own feelings and motivations onto others.
---	-.33a	Feels cheated and victimized by life.
---	.28b	Is physically attractive/good-looking.
-.29b	-.46a	Has fluctuating moods.
---	.37a	Is cheerful.
---	.30b	Handles anxiety/conflicts by repression or disassociations.
-.23c	-.36a	Expresses hostile feelings directly.
-.25b	---	Values own independence and autonomy.
---	.27b	Is emotionally bland/has flattened effect.

Table 4 (Continued)

JHS r	SHS r	Q-Item Content
-.23c	---	Is self-dramatizing/histrionic.
---	.30b	Does not vary role/relates to everyone in the same way.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 5

Personality Rating Correlates in Adulthood of the WILTD Over-Control Score Gathered When the Male Subjects Were in Senior High School

Adult r	Q-Item Content
.40b	Is a genuinely dependable and responsible person.
.38b	Is fastidious.
.45a	Prides self on being objective, rational.
.52a	Tends toward overcontrol of needs and impulses.
-.36b	Thinks and associates ideas unusually.
-.38b	Is facially and/or gesturally expressive.
-.39b	Is unpredictable and changeable in behavior and attitudes.
-.42b	Undercontrol of needs, impulses.
-.30c	Is an interesting, arresting person.
-.48a	Enjoys sensuous experiences.
-.35b	Tends to be rebellious and non-conforming.
-.35b	Pushes/stretches limits/sees what he can get away with
.32c	Has high aspiration level for self.
-.41b	Perceives different contexts in sexual terms.
.42b	Has clear-cut, internally consistent personality.
.36b	Is physically attractive/good-looking
-.32c	Has fluctuating moods.
.43b	Handles anxiety/conflicts by repression or disassociation
.32c	Is power oriented/values power in self and others.
-.31c	Expresses hostile feelings directly.
-.41b	Is self-dramatizing/histrionic.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 6

Personality Rating Correlates in Adulthood of the WILTD Over-Control Score Gathered When the Female Subjects Were in Senior High School

Adult r	Q-Item Content
-.51a	Has a wide range of interests.
-.38c	Appears to have a high degree of intellectual capacity.
.37c	Is uncomfortable with uncertainty and complexities.
.38c	Tends to be self-defensive.
.35c	Extrapunitive/tends to transfer or project blame.
.44b	Is negativistic/tends to undermine/obstruct/sabotage.
.43b	Is moralistic.
.34c	Is self-defeating.
.41b	Responds to humor.
-.58a	Has insight into own motives and behavior.
-.33c	Socially perceptive of wide range of interpersonal cues.
-.35c	Appears straightforward, forthright.
-.37c	Is verbally fluent/can express ideas well.
.48b	Does not vary role/relates to everyone in same way.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level, if followed by a c, at the .10 level. A total of 90 Q-items were evaluated for significance.

Table 7

CCQ-correlates of Gift Delay Time in 3-Year Old Boys and in Girls

Correlation in sample of boys	Correlation in sample girls	CCQ-Item
.19	.26b	Is considerate of other children.
-.34b	-.09	Seeks physical contact with others.
.32b	.20	Tends to keep thoughts and feelings to self.
.08	.29b	Develops genuine and close relationships.
-.19	-.37a	Has transient interpersonal relationships.
-.31b	-.43a	Attempts to transfer blame to others.
.35b	.08	Shows concern for moral issues.
-.23	-.26b	Expresses negative feelings directly and openly.
-.37a	-.25c	Tries to take advantage of others.
.47a	.12	Uses and responds to reason.
-.49a	.12	Is visibly deviant from peers.
.32b	.21	Is protective of others.
.35b	.08	Shows a recognition of others' feelings; empathic.
-.28b	-.05	Cries easily.
-.38a	-.14	Is restless and fidgety.
.27b	.15	Is inhibited and constricted.
-.08	.26b	Is resourceful in initiating activities.
.35b	.07	Tends to withdraw or disengage self under stress.
-.30b	.01	Tends to go to pieces under stress.
-.42a	-.23c	Has rapid shifts in mood; emotionally labile.
-.38a	.20	Is afraid of being deprived; concerned about getting enough.

Table 7 (Continued)

Correlation in sample of boys	Correlation in sample girls	CCQ-Item
-.36a	-.05	Is jealous and envious of others.
.19	-.40a	Tends to dramatize or exaggerate mishaps.
-.09	-.36a	Tends to be judgmental of others' behavior.
-.33b	-.13	Has a rapid personal tempo.
-.43a	-.50a	Is unable to delay gratification.
.35b	.26b	Is attentive and able to concentrate.
.39a	.34b	Is planful, thinks ahead.
.01	.30b	Daydreams, tends to get lost in reverie.
.22	.40a	Becomes strongly involved in what (s)he does.
-.30b	.03	Is a talkative child.
-.30b	-.34a	Is aggressive (physically or verbally).
-.23	-.26b	Is stubborn.
-.36a	-.15	Emotional reactions are inappropriate.
-.18	-.33b	Overreacts to minor frustrations; easily irritated.
.34b	.15	Has an active fantasy life.
.33b	.21	Is shy and reserved; makes social contadts slowly.
.37a	.33b	Is reflective; thinks and deliberates before acting.

Note: Correlations followed by an a are significant at the .01 level; if followed by a b, at the .05 level; if followed by a c, at the .10 level. A total of 100 CCQ items were evaluated for significance.

Footnotes

1. A better way to evaluate the existence of a general or consistency factor is to evaluate the size of the first unrotated factor extracted. Some interpretative caution is still required, however, because whereas the varimax procedure fundamentally destroys a general factor that might be present, the amount of variance explained by the first unrotated factor is slightly overstated because of error fitting.

2. What I call O-data in this essay Cattell labelled L- (for life) data. I prefer the O-designation because it is a continual reminder that this data domain depends quintessentially on the use of an observer as an active, filtering, cumulating, weighting, integrating instrument.

3. What I call S-data Cattell labelled as Q-data when making these data distinctions. I prefer to use the letter S (for self-reporting), as a more general tag for this kind of data and also to avoid the confusion that would arise because the letter Q, before Cattell, had been preempted by Stephenson (1953) to identify the ipsative approach (e.g., as in Q-sorting).

4. I report primarily my own research because it is easiest for me to do and because the task of finding and evaluating a

goodly portion of the relevant evidence in an unorganized literature is beyond my energies and particular interest. I would claim, however, that because of my preoccupation with some of the issues surrounding the consistency-specificity controversy, my research has been sensitive to past concerns and has tried to respond to them. Other psychologists will be able to cite other research that also deserves mention in support of the arguments I am collecting and advancing here. This essay should by no means be considered a survey of extant evidence.

5. Appendices E, F, and G of Lives Through Time (Block, 1971) contain in detail the data here being summarized; Chapter 5 places this information into a psychological context.

6. The earlier CPI protocols were collected and developed for Lives Through Time (Block, 1971); the latest CPI procotols were collected in a subsequent followup assessment of the subjects conducted by the Institute of Human Development. The correlations between the early and later CPI protocols were computed under the auspices of Dr. Jane Brooks who will be reporting in detail on her analyses and their implications. I am most grateful to her for permission to report these data in summary form.

7. The reader will recall that the CPI was given on two separate occasions separated by about ten years. Logically, we would

expect that if CPI scores from one time relate well to early O-data, then CPI scores from the second administration also should relate to the early O-data. And such is the case. Dr. Jane Brooks will be reporting these findings, based on a period approaching 35 years!